Body size and risk for colorectal cancers showing BRAF mutations or microsatellite instability: a pooled analysis

Published by Oxford University Press on behalf of the International Epidemiological Association
International Journal of Epidemiology 2012;**41**:1060–1072
doi:10.1093/ije/dys055

CANCER

Body size and risk for colorectal cancers showing *BRAF* mutations or microsatellite instability: a pooled analysis

Laura AE Hughes,[1] Elizabeth J Williamson,[2] Manon van Engeland,[3] Mark A Jenkins,[2] Graham G Giles,[2,4] John L Hopper,[2] Melissa C Southey,[5] Joanne P Young,[6,7] Daniel D Buchanan,[6] Michael D Walsh,[6] Piet A van den Brandt,[1] R Alexandra Goldbohm,[8] Matty P Weijenberg[1]* and Dallas R English[2,4]

[1]Department of Epidemiology, GROW-School for Oncology and Developmental Biology, Maastricht University, Maastricht, The Netherlands, [2]Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, School of Population Health, University of Melbourne, Melbourne, Victoria, Australia, [3]Department of Pathology, GROW-School for Oncology and Developmental Biology, Maastricht University Medical Center, The Netherlands, [4]Cancer Epidemiology Centre, The Cancer Council Victoria, Carlton, Victoria, Australia, [5]Genetic Epidemiology Laboratory, Department of Pathology, University of Melbourne, Melbourne, Victoria, Australia, [6]Familial Cancer Laboratory, Queensland Institute of Medical Research, Herston, Australia, [7]Faculty of Health Sciences, University of Queensland, Brisbane, Queensland, Australia and [8]Department of Prevention and Health, TNO Quality of Life, Leiden, The Netherlands

*Corresponding author. Department of Epidemiology, Maastricht University, P.O. Box 616, 6200 MD, The Netherlands. E-mail: mp.weijenberg@maastrichtuniversity.nl

Accepted 15 March 2012

Background How body size influences risk of molecular subtypes of colorectal cancer (CRC) is unclear. We investigated whether measures of anthropometry differentially influence risk of tumours according to *BRAF* c.1799T>A p.V600E mutation (*BRAF*) and microsatellite instability (MSI) status.

Methods Data from The Netherlands Cohort Study ($n = 120\,852$) and Melbourne Collaborative Cohort Study ($n = 40\,514$) were pooled and included 734 and 717 colorectal cancer cases from each study, respectively. Hazard ratios (HRs) and 95% confidence intervals (CIs) for body mass index (BMI), waist measurement and height were calculated and compared for subtypes defined by *BRAF* mutation and MSI status, measured from archival tissue.

Results Results were consistent between studies. When pooled, BMI modelled in 5 kg/m^2 increments was positively associated with *BRAF* wild-type (HR: 1.16, 95% CI: 1.08–1.26) and MS-stable tumours (HR: 1.15, 95% CI: 1.06–1.24). Waist measurement was also associated with *BRAF* wild-type (highest vs lowest quartile, HR: 1.59, 95% CI: 1.33–1.90) and MS-stable tumours (highest vs lowest quartile HR: 1.68, 95% CI: 1.31–2.15). The HRs for *BRAF* mutation tumours and MSI tumours were smaller and non-significant, but differences between the HRs by tumour subtypes were not significant. Height, modelled per 5-cm increase, was positively associated with *BRAF* wild-type and *BRAF* mutation tumours, but the HR was greater for tumours with a *BRAF* mutation than *BRAF* wild-type (HR: 1.23, 95% CI: 1.11–1.37, $P_{\text{heterogeneity}} = 0.03$). Similar associations were observed with respect to height and MSI tumours (HR: 1.26, 95% CI: 1.13–1.40, $P_{\text{heterogeneity}} = 0.02$).

Conclusions Generally, overweight increases the risk of CRC. Taller individuals have an increased risk of developing a tumour with a *BRAF* mutation or MSI.

Keywords Colorectal neoplasms, BRAF, microsatellite instability, body mass index, waist circumference, height, cohort study

Introduction

Although it is well documented that body size influences colorectal cancer (CRC) risk,[1] how this risk differs according to molecular phenotype of tumours is less clear. Elucidating such differences may lead to a better understanding of CRC aetiology and prevention.

One well-defined subgroup of CRC arises through the (sessile) serrated pathway, via serrated epithelium or hyperplastic polyps.[2] A number of genetic and epigenetic abnormalities contribute to the resulting adenoma/carcinoma. The somatic mutation of the *BRAF* c.1799T>A p.V600E (*BRAF*) proto-oncogene is thought to occur early in disease progression.[3] Microsatellite instability (MSI), which occurs when short, repetitive DNA sequences undergo an increase or decrease in repeat length,[4] is thought to occur later in the pathway.[2] In sporadic CRC, MSI almost always arises through loss of function in the mismatch repair gene *MLH1*, resulting from promoter methylation.[2] An underlying abnormality of tumours harbouring a *BRAF* mutation and/or MSI is the CpG island methylator phenotype (CIMP),[2] which is characterized by numerous promoter CpG island hypermethylated tumour suppressor and DNA repair genes.[5–9] This in turn is associated with transcriptional silencing of gene expression.[10]

Ogino *et al.*[11,12] recently described an emerging field of science now referred to as molecular pathological epidemiology. This is the integration of molecular pathology and epidemiology in order to gain insight into potential mechanisms of disease aetiology and to analyse environmental risk factors and disease trends in large numbers of unselected cases. Only a handful studies have investigated the association between body size and characteristics of the (sessile) serrated pathway, and the majority of those was case–control in design and only considered body mass index (BMI) as a risk factor.[4,13–17] Central adiposity is thought to be a better predictor of CRC than BMI[18] and height reflects early-life environmental and hormonal exposures,[1] therefore, investigating these risk factors in addition to BMI is warranted.

The Netherlands Cohort Study (NLCS) and the Melbourne Collaborative Cohort Study (MCCS) are prospective cohorts in which colorectal tumour tissue samples have been collected and subsequently analysed for molecular characteristics. We pooled data from individual participants from the two studies to produce a more accurate estimate of relative risk.[19,20] The purpose of the present study was to investigate if body size (BMI, indicators of central adiposity and height) differentially influence risk of tumours according to *BRAF* mutation and MSI status.

Materials and Methods

Study populations and design

NLCS

The NLCS was initiated in 1986 to investigate the association between diet and the development of cancer. It includes 58 279 men and 62 573 women aged 55–69 years at baseline who completed a self-administered food frequency questionnaire involving 150 food items as well as questions on dietary habits, lifestyle, health and demographics.[21,22] Municipal registries from throughout The Netherlands were used to constitute an efficient sampling frame.[21–23] The NLCS uses a case–cohort approach for data processing and analysis; case subjects were derived from the entire cohort and the number of person-years at risk for the entire cohort was estimated from a subcohort of 5000 men and women who were randomly sampled from the full cohort at baseline. All subcohort members who reported prevalent cancer (excluding skin cancer) at baseline were excluded from analyses, leaving 4654. Further details of the NLCS design have been described.[21–23]

Incident CRC cases were identified by annual record linkage to nine regional cancer registries and a national pathology database (PALGA).[24] The completeness of cancer follow-up is almost 100%.[25] Paraffin-embedded tumour material from CRC patients was retrieved, as described previously.[26] In total, 734 incident CRC patients were identified from a follow-up period of 7.3 years after baseline (until December 31, 1993), excluding the first 2 years of follow-up, of whom a PALGA report of the lesion as well as sufficient DNA were available.[26]

The study protocol was approved by the Medical Ethics Committees of the University Hospital Maastricht and TNO Nutrition. Tumour material was collected after approval by the ethical review boards of Maastricht University, the National Cancer Registry and PALGA.

MCCS

The MCCS is a prospective cohort study of 17 045 men and 24 469 women, who were between the age of 27 and 75 years at recruitment from 1990 to 1994 (almost all were aged 40–69 years), and includes 5425 migrants from Italy and 4535 from Greece. For this analysis, 202 participants who had a CRC diagnosed before baseline were excluded, leaving a total of 41 312 individuals. Subjects were recruited via the electoral rolls (registration to vote is compulsory for adults in Australia), advertisements and community announcements in local media (e.g. television, radio and newspapers). Comprehensive lists of Italian and Greek surnames were also used to target Southern European migrants listed in the phone books and on electoral rolls. A structured interview schedule was used to obtain information on potential risk factors. Information on current diet was obtained from a dietary questionnaire that contained a 121-item food frequency questionnaire that was developed for the MCCS.[27]

Cases were participants who had a first diagnosis of invasive cancer of the colon or rectum during follow-up to December 31, 2004, identified by linkage to population-based cancer registries in all Australian states. Addresses and vital status of the subjects were determined by record linkage to electoral rolls, the National Death Index, Victorian death records, from electronic phone books and from responses to mailed questionnaires and newsletters. Archival tumour tissue was sought for all primary, histopathologically confirmed adenocarcinomas diagnosed in Victoria. In the MCCS, 717 tumours were available for analysis.

The study protocol was approved by the Cancer Council Victoria's Human Research Ethics Committee. Participants gave written consent for participation and for the investigators to obtain their medical records.

Anthropometry

NLCS

Height (cm) and body weight (kg) were self-reported on the baseline questionnaire. BMI was subsequently calculated. At baseline, individuals were also asked to report their lower body (trouser or skirt) clothing size from their clothing label (Dutch sizes). Clothing size appears to predict cancer risk independently of BMI and corresponds to waist measurements.[28]

MCCS

Height (cm), weight (kg) and waist circumference (cm) were measured at baseline attendance for each participant according to written protocols that were based on standard procedures.[29] Weight was measured to 100 g using digital electronic scales, height to 1 mm using a stadiometer and waist circumference was measured to 1 mm using a 2 -m metal anthropometric tape. BMI was calculated from weight and height.

Molecular pathology

Although CIMP has been measured in both studies, it was not considered here because different methods of measurement were used, leading to different proportions of CRC s showing the phenotype.

NLCS

BRAF mutation analysis was done by a semi-nested PCR and subsequent restriction fragment length polymorphism (RFLP) analyses as previously described.[30] MSI was determined by a pentaplex PCR using the MSI markers BAT-26, BAT-25, NR-21, NR-22 and NR-24, as described in detail by Suraweera *et al*.[31] Tumours were classified as MSI-high if two or more markers showed instability and as MSI-low/MS-stable (MSS) if one or none of the markers examined showed instability.

MCCS

BRAF mutation analysis was done by a real-time PCR-based allelic discrimination method.[32] MSI was examined using 10 microsatellite markers.[33] High-frequency MSI (MSI-H) was defined as >30% instability of at least five markers; low-level MSI (MSI-L) was defined as 1–29% of loci unstable. Tumours were classified as MSI-high, MSI-intermediate, MSI-low and MS-stable, but for comparability with the NLCS, MSI-low and MS-stable were combined into one category.

In both NLCS and MCCS, cases with and without tumour tissue retrieved were similar with respect to established risk factors for CRC, indicating that no selection bias was introduced as a result of inability to retrieval tissue for all cases.

Statistical analyses

Data were analysed with Stata (version 10, Statacorp, College Station, TX, USA). Hazard ratios (HRs), 95% confidence intervals (95% CIs) and *P*-values were calculated using Cox regression with age as the time-scale.[34] To account for its case–cohort design, survival times in the NLCS were adapted using Prentice weights.[35]

To estimate HRs separately for different molecular subtypes and to test their difference (i.e. *BRAF* mutation vs *BRAF* wild-type), Cox models based on competing risks were fitted using a data duplication method.[36] Person-years of follow-up were calculated from baseline until the date of CRC diagnosis, death or end of follow-up.

Study-specific analyses

BMI (per 5 kg/m^2) and height (per 5 cm) were fitted as continuous covariates to estimate linear trends on the log hazard scale and as categorical variables. These variables were additionally considered as sex- and study-specific quartiles. With respect to waist measurements, trouser/skirt size and waist

circumference were fitted categorically for NLCS and MCCS data, respectively, according to approximate sex-specific quartiles. All models were stratified by sex and the MCCS was additionally stratified by ethnicity (based on country of birth and classified into two groups: Southern European (born in Italy or Greece) and Anglo-Celtic (born in Australia, New Zealand, UK or Ireland).[37]

Several potential confounding variables were considered for multivariate analysis. For the NLCS, these included family history of CRC (yes/no), smoking status (never smoker, ex-smoker and current smoker), education (primary school, junior high school, senior high school, higher vocational school or university), total energy intake (kcal/day), alcohol intake (0, 0.1–4, 5–14, 15–29 and $\geqslant$30 g/day), recreational physical activity (<30, 30–60, 61–90 and >90 min/day) and consumption of red meat, fruit, vegetables, fibre and grains (g/day). For women, hormonal factors such as contraceptive use (yes/no), HRT use, age at menarche, age at menopause and number of children were also tested. For the MCCS, the variables considered were smoking status (current smoker, former smoker and never smoker), education (primary school, some high/technical school, completed high school and completed tertiary degree/ diploma), total dietary energy intake (kcal/day), current alcohol consumption (males 0, 1–39 and $\geqslant$40 g/day; females 0, 1–19 and $\geqslant$20 g/day), current level of physical activity (none, low, moderate and high), meat, fruit and vegetable (servings/day) intake, fibre from cereal products, multivitamin and fibre supplements (yes/no). For women, parity, total months of lactation for all live births, age at first live birth, age at menarche, HRT use (never, former and current), oral contraceptive pill usage (never and ever), menopausal status and age at menopause were also considered. In the respective studies, no variables changed the estimated HRs by >10% and therefore none was included in the final models.

Pooled analyses

Models were stratified by study, ethnicity and sex. HRs, 95% CIs and *P*-values were estimated for a 5-kg/m^2 increase in BMI and a 5-cm increase in height. Waist measurement was not considered continuously for the pooled data, because of differences in measurement. BMI, waist measurement and height were additionally considered according to sex and study-specific quartiles. The NLCS was coded to reflect a non-Southern European ethnicity. Due to the increase in sample size, sex-specific analyses were also conducted.

Tests based on Schoenfeld residuals and graphical methods using Kaplan–Meier curves showed no evidence that proportional hazard assumptions were violated for any analyses.

Results

Generally, men and women in the NLCS were older, taller and had a lower BMI than those in the MCCS. There was a higher proportion of current and ex-smokers in the NLCS than in the MCCS (Tables 1 and 2).

Of the 734 incident tumours in the NLCS, 697 were successfully analysed for *BRAF* p.V600E c.1799T>A mutation status and 658 for MSI (Table 3). Of the 717 tumours identified in the MCCS, *BRAF* mutation status was determined for 582 and MSI status for 585. Further details of the cancers in each cohort are in previous publications.[37–40] For both studies, 16% of the cancers evaluated had *BRAF* mutations. In the NLCS, 14% of tumours from males and 19% from females had *BRAF* mutations, whereas in the MCCS, the proportions were 11 and 19%, respectively. Cases with *BRAF* mutation tended to be older and taller than non-*BRAF* cases. With respect to MSI status, 13% of the cancers evaluated in the NLCS and 15% in the MCCS were MSI high. In the NLCS, 12% of the tumours from males and 14% from females were MSI, whereas in the MCCS, the proportions were 12 and 18%, respectively. MSI-high cases tended to be taller than MSS cases

Associations between body size and CRC risk according to *BRAF* status are shown in Table 4. For both studies, BMI modelled per 5-kg/m^2 increase was associated with *BRAF* wild-type tumours (NLCS HR: 1.28/5 kg/m^2, 95% CI: 1.12–1.45; MCCS HR: 1.10, 95% CI: 0.99–1.21). This association remained when the data were pooled (HR: 1.16, 95% CI: 1.08–1.26). After pooling and considering BMI in sex and study-specific quartiles, there was a positive dose–response association only for *BRAF* wild-type tumours and no association with respect to tumours with a *BRAF* mutation. Tests for heterogeneity between *BRAF* subtypes were not statistically significant.

With respect to waist measurements in the MCCS, there were similar dose–response associations for both *BRAF*-mutated and *BRAF* wild-type tumours when considering quartiles of waist circumference; however, these only reached statistical significance for the *BRAF* wild-type group. This was also observed in the NLCS for clothing size, although there was less evidence of a dose–response. When the data were pooled, there was a statistically significant positive dose–response association for *BRAF* wild-type tumours (highest vs lowest quartile HR: 1.59, 95% CI: 1.33–1.90 and P trend < 0.001). A statistically significant trend was also observed for *BRAF*-mutated tumours ($P = 0.03$).

When height was modelled per 5-cm increment, there was an increased risk for both tumours with and without *BRAF* mutations, although the HR was greater for *BRAF*-mutated tumours. For the pooled data, the HR for tumours with *BRAF* mutations was 1.23/5-cm increment (95% CI: 1.11–1.37) compared with 1.08 (95% CI: 1.03–1.13) for tumours without

Table 1 Baseline demographic, anthropometric, dietary and lifestyle characteristics of NLCS

Characteristics	Males	Females
Total in cohort	58 279	62 573
Total in subcohort[a]	2232	2399
Age (years)[b]	61.3 (4.2)	61.5 (4.3)
Weight (kg)	77.8 (9.5)	68.5 (10.3)
Height (cm)	176.4 (6.7)	165.1 (6.3)
BMI (kg/m^2)	25.0 (2.6)	25.1 (3.6)
Trouser size (Dutch size)	51 (4)	44 (3)
Smoking (%)		
Former	51	20
Current	36	21
Never	13	59
Alcohol intake (%)		
0 g/day	14	33
0.1–4	21	36
5–14	28	19
15–29	23	9
$\geqslant$30	15	3
Education (%)		
Primary school	27	35
Some high school	21	23
Completed high school	34	33
Completed tertiary school	18	8
Family history of CRC (%)		
No	95	94
Yes	5	6
Recreational physical activity (%)		
$<$30 min/day	19	27
30–60	32	31
60–90	18	22
$\geqslant$90	31	21
Total energy (kcal/day)	2148 (526)	1655 (417)
Red meat (g/day)	104.8 (44.2)	92.3 (41.2)
Fibre (g/day)	28.4 (9.0)	24.9 (7.3)
Folage (μg/day)	208.5 (76.3)	195.2 (70.9
Hormone replacement therapy		
Ever	–	12
Never	–	85
Unknown	–	3

[a]The NLCS uses a case–cohort approach for data processing and estimating person-time at risk for the entire cohort. This subcohort consists of 5000 individuals, of whom 4631 were available for the present analysis.
[b]Mean (SD) or percentages where indicated.

Table 2 Baseline demographic, anthropometric, dietary and lifestyle characteristics of the MCCS

Characteristics	Males	Females
Total in cohort	16 942	24 370
Age (years)[a]	55.8 (8.8)	55.0 (8.6)
Weight (kg)	80.8 (11.8)	68.2 (12.4)
Height (cm)	172.4 (7.4)	159.8 (6.7)
BMI (kg/m^2)	27.2 (3.6)	26.7 (4.9)
Waist circumference (cm)	93.5 (10.0)	80.0 (11.8)
Smoking (%)		
Former	45	22
Current	15	9
Never	41	69
Alcohol intake (%)[b]		
Lifetime abstainer	14	39
Ex-drinker	5	3
Low intake	27	19
Medium intake	29	21
High intake	25	17
Education (%)		
Primary school	19	20
Some high school	31	43
Completed high school	25	18
Completed tertiary school	25	19
Recreational physical activity (%)		
None	23	22
Low	18	21
Moderate	34	36
High	25	20
Total energy (kcal/day)	2468 (777)	2023 (666)
Red meat (times/week)		
0–4	18	28
5–6	21	25
7–9	27	27
$\geqslant 10$	34	21
Fibre (g/day)	32.2 (12.7)	30.0 (11.2)
Folate (μg/day)	330.2 (149.0)	321.3 (143.9)
Hormone replacement therapy (%)		
Never	–	75
Former	–	9
Current	–	17

[a]Mean (SD) or percentages where indicated.
[b]Sex-specific categorization.

Table 3 Characteristics of participants in the NLCS and MCCS by CRC status

Characteristics	No cancer	CRC diagnosed during follow-up			
		BRAF V600E mutation status		MSI status	
		Mutation	Wild-type	MSI-high	MSI-stable
NLCS					
Total (%)	4631[a]	112 (16)[b]	585 (84)	84 (13)	574 (87)
Sex (%)					
Male	48	47	57	50	55
Female	52	53	43	50	45
Age (years)[c]	61.4 (4.2)	62.8 (4.1)	63.0 (4.1)	62.9 (4.0)	63.5 (4.5)
Height (cm)					
Male	176.4 (6.7)	177.8 (7.9)	176.6 (6.6)	177.0 (7.3)	176.9 (6.8)
Female	165.1 (6.2)	166.5 (5.3)	166.0 (6.6)	167.2 (5.3)	166.1 (6.6)
BMI (kg/m^2)	25.1 (3.1)	25.4 (2.9)	25.6 (3.2)	25.6 (3.3)	25.5 (3.1)
Clothing size					
Male	51 (4)	52 (4)	52 (4)	52 (2)	52 (3)
Female	44 (3)	44 (3)	44(3)	44 (3)	44 (3)
MCCS					
Total (%)	40 595	95 (16)	487 (84)	90 (15)	495 (85)
Sex (%)					
Male	41	35	53	40	52
Female	59	65	47	60	48
Age (years)	55.2 (8.7)	62.4 (6.4)	60.4 (7.4)	60.4 (7.9)	60.7 (7.3)
Height (cm)					
Male	172.4 (7.4)	173.8 (6.5)	171.8 (7.0)	176.1 (5.9)	171.4 (6.9)
Female	159.8 (6.7)	161.4 (6.6)	159.4 (6.7)	159.7 (6.0)	159.9 (7.0)
BMI (kg/m^2)	26.9 (4.4)	26.8 (4.3)	27.5 (4.1)	26.9 (4.1)	27.5 (4.2)
Waist circumference (cm)					
Male	93.4 (10.0)	95.6 (9.5)	96.2 (9.4)	94.1 (8.9)	96.5 (9.4)
Female	79.9 (11.8)	82.7 (11.8)	82.6 (12.3)	83.1 (11.0)	82.6 (12.5)
Pooled data					
Total (%)	45 226	207 (16)	1072 (84)	174 (14)	1069 (86)
Sex (%)					
Male	42	42	55	45	54
Female	58	58	45	55	46
Age	55.8 (8.5)	62.6 (5.2)	61.8 (6.0)	61.9 (6.7)	61.9 (5.9)
Height					
Male	172.9 (7.4)	176.3 (7.6)	174.5 (7.2)	176.6 (6.7)	174.4 (7.3)
Female	160.3 (6.8)	163.8 (6.5)	162.8 (7.4)	162.9 (6.8)	163.0 (7.4)
BMI	26.7 (4.4)	26.0 (3.7)	26.5 (3.8)	26.3 (3.7)	26.4 (3.8)
Site (%)					
Proximal colon	–	62	27	74	27
Distal colon	–	17	31	13	30
Rectosigmoid/rectum	–	18	39	9	40
Unknown	–	2	3	4	3
MSI status (%)					
MSI-stable	–	57	92	–	–
MSI-high	–	43	8	–	–

[a]Subcohort derived from total cohort to estimate person-time at risk.
[b]*N* (%).
[c]Mean (SD) or percentages where indicated.

Table 4 HRs and 95% CIs for the association between BMI, waist measurements, height and *BRAF* mutation status of colorectal tumours in the NLCS, MCCS and a pooled analysis of the two studies

	NLCS					MCCS					Pooled analysis			
	BRAF mutation			*BRAF* wild-type		*BRAF* mutation			*BRAF* wild-type		*BRAF* mutation		*BRAF* wild-type	
	PY	*N*	HR (95% CI)[a]	*N*	HR (95% CI)[a]	PY	*N*	HR (95% CI)[a]	*N*	HR (95% CI)[a]	*N*	HR (95% CI)[b]	*N*	HR (95% CI)[b]
Continuous														
BMI														
Per 5 kg/m^2	35 129	109	1.11 (0.85–1.45)	550	1.28 (1.12–1.45)	466 772	95	1.01 (0.80–1.27)	487	1.10(0.99 –1.21)	204	1.05 (0.88–1.25)	1037	1.16 (1.08–1.26)
$P_{heterogeneity}$[c]		0.35				0.50					0.29			
Height (cm)														
Per 5 cm	34 351	110	1.19 (1.03–1.37)	561	1.08 (1.01–1.15)	466 804	95	1.29 (1.10–1.52)	487	1.07 (1.00–1.15)	204	1.23 (1.11–1.37)	1044	1.08 (1.03–1.13)
$P_{heterogeneity}$		0.23				0.04					0.03			
Quartiles[d]														
BMI (kg/m^2)														
1	11 133	28	1 (referent)	140	1 (referent)	114 954	25	1 (referent)	92	1 (referent)	53	1 (referent)	232	1 (referent)
2	6632	26	1.57 (0.92–2.67)	89	0.99 (0.76–1.29)	116 243	22	0.81 (0.46–1.44)	123	1.22 (0.93–1.60)	48	1.14 (0.77–1.70)	212	1.11 (0.92–1.34)
3	11 445	37	1.24 (0.76–2.02)	204	1.39 (1.12–1.72)	117 680	27	0.99 (0.58–1.71)	128	1.20 (0.91–1.58)	64	1.12 (0.78–1.61)	332	1.31 (1.11–1.55)
4	5917	18	1.22 (0.67–2.21)	117	1.40 (1.09–1.80)	117 895	21	0.87 (0.49–1.55)	144	1.35 (1.03–1.78)	39	1.04 (0.69–1.58)	261	1.38 (1.15–1.66)
P_{trend}			*0.56*		*<0.001*			*0.83*		*0.04*		*0.77*		*<0.001*
Waist measurement														
1	13 725	32	1 (referent)	176	1 (referent)	118 689	17	1 (referent)	80	1 (referent)	49	1 (referent)	256	1 (referent)
2	10 595	36	1.37 (0.85–2.21)	176	1.16 (0.94–1.43)	111 966	21	1.14 (0.61–2.13)	101	1.22 (0.91–1.64)	57	1.28 (0.87–1.87)	277	1.18 (0.99–1.40)
3	6285	24	1.46 (0.86–2.48)	105	1.15 (0.91–1.46)	121 233	28	1.38 (0.77–2.51)	136	1.45 (1.10–1.92)	52	1.42 (0.95–2.11)	241	1.28 (1.07–1.53)
4	3747	11	1.15 (0.58–2.27)	84	1.59 (1.23–2.06)	114 810	29	1.52 (0.84–2.72)	163	1.66 (1.27–2.18)	40	1.40 (0.92–2.13)	274	1.59 (1.33–1.90)
P_{trend}			*0.21*		*0.009*			*0.07*		*<0.001*		*0.03*		*<0.001*
Height (cm)														
1	9954	26	1 (referent)	146	1 (referent)	120 026	18	1 (referent)	135	1 (referent)	44	1 (referent)	281	1 (referent)
2	9049	26	1.12 (0.65–1.92)	141	1.11 (0.88–1.40)	122 910	23	1.13 (0.61–2.09)	141	1.13 (0.88–1.43)	49	1.11 (0.74–1.67)	282	1.12 (0.95–1.32)
3	9150	27	1.17 (0.68–2.01)	143	1.15 (0.92–1.45)	111 235	25	1.45 (0.79–2.66)	98	0.95 (0.72–1.25)	52	1.28 (0.86–1.92)	241	1.06 (0.89–1.26)
4	7195	30	1.73 (1.02–2.92)	127	1.29 (1.03–1.64)	112 634	29	2.08 (1.13–3.81)	113	1.31 (1.00–1.73)	59	1.87 (1.26–2.77)	240	1.31 (1.09–1.56)
P_{trend}			*0.03*		*0.04*			*0.05*		*0.17*		*0.004*		*0.01*

[a]Analysis stratified by sex and ethnicity.
[b]Analysis stratified by sex, ethnicity and study.
[c]*P*-value for test that HRs for two tumour subtypes are equal.
[d]Sex- and study-specific quartile boundaries are as follows: BMI—NLCS: men: <23.4, 23.4–24.8, 24.9–26.5, ⩾26.6 kg/m^2 and women: <23.8, 23.8–24.6, 24.7–29.9, ⩾29.9 kg/m^2; MCCS: men <24.8, 24.8–26.8, 26.9–29.1, ⩾29.2 kg/m^2 and women: <23.2, 23.2–25.9, 25.9–29.4, ⩾29.4 kg/m^2; clothing size—NLCS: men <50, 52, 54 ⩾ 56; women <42, 44, 46 ⩾48; waist circumference—MCCS: men <87.0, 87.0–93.0, 93.1–99.5 ⩾99.5 cm; women <71.2, 71.3–78.3, 78.3–87.0 ⩾87.1 cm; height—NLCS: men <173, 173–176, 177–182, ⩾183 cm and women <162, 162–165, 166–171 ⩾172 cm; MCCS men <167.6, 167.6–172.5, 172.6–177.5, ⩾177.6 cm and women <155.3, 155.3–160.0, 160.0–164.3, ⩾164.4 cm.

BRAF mutations ($P_{heterogeneity} = 0.03$). Height was also associated with both tumour subtypes when modelled according to sex- and study-specific quartiles, showing a positive dose–response association that was stronger for mutated than wild-type tumours.

Table 5 shows associations according to MSI status. HRs for pooled BMI and waist measurements were similar to those observed for *BRAF*; a statistically significant association was observed only for MSS tumours. Height was also associated with both tumour subtypes. For the pooled data, the HR for MSI-high tumours was 1.26/5 cm increase (95% CI: 1.13–1.40) compared with 1.08 (95% CI: 1.03–1.14) for MSS tumours ($P_{heterogeneity} = 0.02$).

Stratifying the analyses according to sex showed no differences between men and women.

Discussion

We present prospective cohort data on associations between body size and molecular subsets of CRC, specifically *BRAF* mutation and MSI status. We pooled data from the NLCS and MCCS and observed that there was little heterogeneity between tumour subtypes when considering associations according to BMI and waist measurement; however, BMI and waist measurements were strong risk factors for tumours without *BRAF* mutations or MSI. Interestingly, associations with height showed more heterogeneity; our findings suggest that height is a stronger risk factor for tumours with *BRAF* mutations or MSI.

This is the first molecular pathological epidemiology study to use pooled data from demographically and operationally diverse cohorts. Both the NLCS and MCCS have almost complete ascertainment of CRC and little loss to follow-up. In the MCCS, body size at baseline was measured according to standard protocols. Measures of body size in the NLCS were obtained by self-report; however, there are many examples in the literature showing that this method is a valid and reliable tool for assessing body weight and height in cohort studies.[41–44] As *BRAF*-mutated tumours and MSI are uncommon, pooling the NLCS and MCCS allows for greater precision than in the individual studies. Meta-analysis is also a strategy that may be considered when combining the data of studies. With a meta-analysis, analysis can be performed independently at two sites and is therefore more economical and easier to perform than pooling of data. However, an important additional feature of a pooled analysis is that it is carried out on individual data. This means that comparability between two studies can be increased, as study subjects can be reclassified with respect to exposure, outcome and confounders.[45] A case study comparing the outcome of the two approaches suggests that although the approaches differ in the statistical analysis, conclusions reached from meta-analysis and pooled analysis are broadly consistent.[45]

To our knowledge, only one study has investigated the association between BMI and *BRAF* status in CRC tumours. In a case–control study, Slattery *et al.*[14] reported that obesity (BMI $\geqslant 30\,kg/m^2$) was not associated with *BRAF*-mutated tumours (associations with *BRAF* wild-type tumours were not reported). Central adiposity has been identified as a stronger predictor of CRC than BMI[1,18] and based on our observations, we propose that waist measurements may be a better predictor of CRC than BMI when analysing data according to molecular subtype, especially when there is limited power to detect associations. Furthermore, it may reduce heterogeneity between studies. When HRs were estimated according to waist measurements, the associations observed in the study-specific analyses were more comparable with each other than those observed for BMI and the pooled analysis suggested a dose–response association with respect to both *BRAF* wild-type and MSS tumours. Future studies should consider this variable in addition to BMI.

Two case–control studies have previously considered the association between BMI and MSI.[4,46] Whereas the study by Campbell *et al*.[4] distinguished between MSS and MSI-low tumours, we combined these two groups because our MSI-low group was very small. The presence and role of MSI-low remains controversial.[47] Mutations often observed in MSI-high tumours appear to be absent in MSI-low tumours and no large differences in clinical or molecular characteristics have been observed between MSS and MSI-low patients.[48] Campbell *et al.* reported that BMI was associated only with MSS and MSI-low tumours, but not MSI-high tumours. We also observed a positive association between BMI and MSS tumours; however, tests for heterogeneity between tumour subtypes did not reach statistical significance in our study. Slattery *et al.*[46] reported associations according to MSI-positive (equivalent to MSI high) or MSI-negative (equivalent to MSS/MSI-low) status, with findings similar to that of Campbell *et al*. and additionally reported a positive association between BMI and MSI-positive tumours in women. When we performed sex-specific analyses on the pooled data, we observed no differences between men and women and no associations with MSI tumours. Neither case–control study reported associations according to waist circumference.

A common feature of tumours characterized by *BRAF* mutations and MSI is CIMP. Although CIMP was measured in both studies, it was not considered here because different methods of measurement were used, leading to different proportions of CRCs showing the phenotype. This highlights a unique challenge of pooling molecular data, as well as the fact that the use of CIMP as an indicator of epigenetic instability has not been without controversy.[49] The association between body size and CIMP has been examined in the NLCS cohort.[16] Although evidence

Table 5 HRs and 95% CIs for the association between BMI, waist measurements, height and MSI status of colorectal tumours in the NLCS, MCCS and a pooled analysis of the two studies

	NLCS					MCCS					Pooled analysis				
	MSI-high			MS-stable		MSI-high			MS-stable		MSI-high		MS-stable		
	PY	*N*	HR (95% CI)[a]	*N*	HR (95% CI)[a]	PY	*N*	HR (95% CI)[a]	*N*	HR (95% CI)[a]	*N*	HR (95% CI)[b]	*N*	HR (95% CI)[b]	
Continuous															
BMI															
Per 5 kg/m^2	35 129	81	1.27 (0.92–1.77)	544	1.22 (1.08–1.39)	466 772	90	0.98 (0.78–1.24)	495	1.10 (1.00–1.22)	171	1.07 (0.89–1.30)	1039	1.15 (1.06–1.24)	
$P_{heterogeneity}$[c]		0.81				0.36					0.53				
Height (cm)															
Per 5 cm	34 351	83	1.20 (1.03–1.41)	548	1.10 (1.03–1.17)	466 804	90	1.32 (1.13–1.54)	495	1.07 (0.99–1.15)	171	1.26 (1.13–1.40)	1043	1.08 (1.03–1.14)	
$P_{heterogeneity}$		0.29				0.02					0.02				
Quartiles[d]															
BMI (kg/m^2)															
1	11 133	23	1 (referent)	140	1 (referent)	114 954	19	1 (referent)	98	1 (referent)	42	1 (referent)	238	1 (referent)	
2	6632	14	1.02 (0.53–2.00)	97	1.08 (0.83–1.41)	116 243	27	1.30 (0.72–1.34)	118	1.10 (0.84–1.43)	41	1.19 (0.77–1.84)	215	1.10 (0.91–1.32)	
3	11 445	29	1.20 (0.69–2.07)	198	1.35 (1.09–1.67)	117 680	23	1.07 (0.57–2.00)	133	1.18 (0.90–1.54)	52	1.14 (0.76–1.72)	331	1.27 (1.08–1.51)	
4	5917	15	1.21 (0.62–2.37)	109	1.33 (1.03–1.71)	117 895	21	1.04 (0.55–1.96)	146	1.31 (1.01–1.71)	36	1.11 (0.70–1.76)	255	1.33 (1.11–1.60)	
P_{trend}			0.46		0.04			0.90		0.04		0.67		<0.001	
Waist measurement															
1	13 725	31	1 (referent)	160	1 (referent)	118 689	17	1 (referent)	80	1 (referent)	39	1 (referent)	255	1 (referent)	
2	10 595	33	1.19 (0.69–2.06)	166	1.20 (0.97–1.48)	111 966	21	1.61 (0.81–3.20)	101	1.14 (0.86–1.53)	48	1.33 (1.87–2.03)	276	1.18 (0.99–1.39)	
3	6285	23	0.88 (0.44–1.79)	100	1.26 (0.99–1.59)	121 233	28	1.97 (1.01–3.82)	136	1.37 (1.04–1.80)	42	1.37 (0.88–2.13)	247	1.30 (1.09–1.56)	
4	3747	10	1.47 (0.72–3.00)	78	1.51 (1.16–1.97)	114 810	29	1.59 (0.80–3.15)	163	1.66 (1.28–2.16)	35	1.40 (0.87–2.24)	248	1.60 (1.33–1.91)	
P_{trend}			0.74		0.04			0.43		<0.001		0.43		<0.001	
Height (cm)															
1	9954	15	1 (referent)	147	1 (referent)	120 026	18	1 (referent)	136	1 (referent)	33	1 (referent)	283	1 (referent)	
2	9049	26	2.00 (1.06–3.76)	126	0.98 (0.77–1.24)	122 910	18	1.02 (0.53–1.95)	147	1.14 (0.90–1.45)	44	1.44 (0.92–2.25)	273	1.06 (0.90–1.26)	
3	9150	22	1.73 (0.90–3.33)	137	1.08 (0.86–1.37)	111 235	27	1.88 (1.02–3.46)	97	0.91 (0.69–1.19)	49	1.79 (1.15–2.80)	234	1.01 (0.84–1.20)	
4	7195	20	2.07 (1.06–4.06)	138	1.38 (1.09–1.73)	112 634	27	2.29 (1.21–4.35)	115	1.30 (0.99–1.70)	47	2.18 (1.38–3.44)	253	1.35 (1.13–1.60)	
P_{trend}			0.20		0.04			0.003		0.38		0.003		0.006	

[a]Analysis stratified by sex and ethnicity.
[b]Analysis stratified by sex, ethnicity and study.
[c]*P*-value for test that HRs for two tumour subtypes are equal.
[d]Sex- and study-specific quartile boundaries are as follows: BMI—NLCS: men <23.4, 23.4–24.8, 24.9–26.5, ⩾26.6 kg/m^2 and women <23.8, 23.8–24.6, 24.7–29.9, ⩾29.9 kg/m^2; MCCS: men <24.8, 24.8–26.8, 26.9–29.1, ⩾29.2 kg/m^2 and women <23.2, 23.2–25.9, 25.9–29.4, ⩾29.4 kg/m^2; clothing size—NLCS: men <50, 52, 54 ⩾ 56; women <42, 44, 46 ⩾48; waist circumference—MCCS: men <87.0, 87.0–93.0, 93.1–99.5 ⩾99.5 cm; women <71.2, 71.3–78.3, 78.3–87.0 ⩾87.1 cm; height—NLCS: men <173, 173–176, 177–182, ⩾183 cm and women <162, 162–165, 166–171 ⩾172 cm; MCCS: men <167.6, 167.6–172.5, 172.6–177.5, ⩾177.6 cm and women <155.3, 155.3–160.0, 160.0–164.3, ⩾164.4 cm.

now suggests that CIMP should be discussed within the context of CIMP-high, CIMP-low and CIMP-negative tumours,[50–52] in the NLCS, associations were considered according to CIMP-high and CIMP-negative tumours and those data suggested that a large body size increases the risk of both subtypes. This supports the observations reported in the present article.

Our observations from the pooled data are intriguing. For all molecular subtypes considered, we observed positive associations, which is consistent with substantial evidence that adult attained height is a strong risk factor for CRC in general.[1] However, HRs were significantly higher for tumours with *BRAF* mutations and MSI than for *BRAF* wild-type or MSS tumours, respectively. Height is a marker of an aggregated fetal and childhood experience and can be considered a proxy measure for important nutritional exposures, which affect several hormonal and metabolic axes.[1] Several studies show that childhood energy restriction is associated with a decreased risk of CRC later in life.[53–57] Furthermore, we recently reported from the NLCS that exposure to severe energy restriction during childhood and adolescence was associated with a low risk of developing a CIMP-positive tumour.[58] Although little research has been done in this area, our current findings contribute to the hypothesis that genetic and epigenetic events in CRC development may be influenced by early-life environmental exposures. This hypothesis may explain why there was little heterogeneity between tumour subtypes when considering associations according to BMI and waist circumferences; these were measured at baseline in adulthood. Analysing additional cohorts to draw firmer conclusions is necessary.

Conclusion

In conclusion, our findings reiterate the importance of a healthy body weight in CRC prevention. Furthermore, positive associations between height, *BRAF* mutation and MSI provide more evidence for the hypothesis that early life events may influence genetic and epigenetic mechanisms.

Funding

World Cancer Research Fund (WCRF) (Grant No. 2007/54 to M.P.W.); European Association for Cancer Research (travel fellowship to L.A.E.H.); Dutch Cancer Society (KWF) (travel fellowship to L.A.E.H.).

Acknowledgements

Special thanks to Julie Simpson for her statistical advice on pooling the NLCS and MCCS data. With respect to the NLCS, we are indebted to the participants of the study and thank the cancer registries (IKA, IKL, IKMN, IKN, IKO, IKR, IKST, IKW, IKZ and VIKC) and The Netherlands Nationwide Registry Pathology (PALGA). We also thank S van de Crommert and J Nelissen for their assistance with data entry and data management; A Volovics and A Kester for statistical advice; L Schouten, C de Zwart, M Moll, W van Dijk, H Brants and A Pisters for assistance; and H van Montfort, T van Moergastel, L van den Bosch and R Schmeitz for programming assistance. With respect to the MCCS, we thank Veronika Gavdik for laboratory assistance, Lesley Jaskowski for the preparation of data for this article and the pathology laboratories that provided archived tissue for the project. This study was made possible by the contribution of many people, including the original investigators, the Program Manager, Georgina Marr, and the diligent team who recruited the participants and who continue working on follow-up. The authors express their gratitude to the many thousands of Melbourne residents who continue to participate in the study.

Conflict of interest: None declared.

KEY MESSAGES

- This study shows strengths and weaknesses of pooling molecular pathological epidemiological data.
- We observed that overweight increases the risk of CRC, regardless of the molecular phenotypes investigated.
- Taller individuals have an increased risk of developing a tumour with a *BRAF* mutation or MSI.

References

[1] World Cancer Research Fund/American Institute for Cancer Research. *Food, Nutrition, Physical Activity, and the Prevention of Cancer: A Global Perspective*. Washington DC: AIRC, 2007.

[2] Noffsinger AE. Serrated polyps and colorectal cancer: new pathway to malignancy. *Annu Rev Pathol* 2009;**4:**343–64.

[3] Jass JR. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. *Histopathology* 2007;**50:**113–30.

[4] Campbell PT, Jacobs ET, Ulrich CM *et al*. Case–control study of overweight, obesity, and colorectal cancer risk, overall and by tumor microsatellite instability status. *J Natl Cancer Inst* 2010;**102:**391–400.

[5] Samowitz WS, Albertsen H, Herrick J *et al*. Evaluation of a large, population-based sample supports a CpG island methylator phenotype in colon cancer. *Gastroenterology* 2005;**129:**837–45.

[6] Shen L, Toyota M, Kondo Y *et al*. Integrated genetic and epigenetic analysis identifies three different subclasses of colon cancer. *Proc Natl Acad Sci U S A* 2007;**104:**18654–59.

[7] Toyota M, Ahuja N, Ohe-Toyota M *et al*. CpG island methylator phenotype in colorectal cancer. *Proc Natl Acad Sci U S A* 1999;**96:**8681–86.

[8] Weisenberger DJ, Siegmund KD, Campan M *et al*. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with *BRAF* mutation in colorectal cancer. *Nat Genet* 2006;**38:**787–93.

[9] Barault L, Charon-Barra C, Jooste V *et al*. Hypermethylator phenotype in sporadic colon cancer: study on a population-based series of 582 cases. *Cancer Res* 2008;**68:**8541–46.

[10] Ferguson LR, Karunasinghe N, Philpott M. Epigenetic events and protection from colon cancer in New Zealand. *Environ Mol Mutagen* 2004;**44:**36–43.

[11] Ogino S, Chan AT, Fuchs CS *et al*. Molecular pathological epidemiology of colorectal neoplasia: an emerging transdisciplinary and interdisciplinary field. *Gut* 2011;**60:**397–411.

[12] Ogino S, Stampfer M. Lifestyle factors and microsatellite instability in colorectal cancer: the evolving field of molecular pathological epidemiology. *J Natl Cancer Inst* 2010; **102:**365–67.

[13] Caan BJ, Coates AO, Slattery ML *et al*. Body size and the risk of colon cancer in a large case–control study. *Int J Obes Relat Metab Disord* 1998;**22:**178–84.

[14] Slattery ML, Curtin K, Sweeney C *et al*. Diet and lifestyle factor associations with CpG island methylator phenotype and *BRAF* mutations in colon cancer. *Int J Cancer* 2007; **120:**656–63.

[15] Slattery ML, Curtin K, Wolff RK *et al*. Diet, physical activity, and body size associations with rectal tumor mutations and epigenetic changes. *Cancer Causes Control* 2010; **21:**1237–45.

[16] Hughes LA, Simons CC, van den Brandt PA *et al*. Body size, physical activity and risk of colorectal cancer with or without the CpG Island Methylator Phenotype (CIMP). *PloS One* 2011;**6:**e18571.

[17] Satia JA, Keku T, Galanko JA *et al*. Diet, lifestyle, and genomic instability in the North Carolina Colon Cancer Study. *Cancer Epidemiol Biomarkers Prev* 2005;**14:**429–36.

[18] Pischon T, Lahmann PH, Boeing H *et al*. Body size and risk of colon and rectal cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC). *J Natl Cancer Inst* 2006;**98:**920–31.

[19] Taioli E, Bonassi S. Pooled analysis of epidemiological studies involving biological markers. *Int J Hyg Environ Health* 2003;**206:**109–15.

[20] Smith-Warner SA, Spiegelman D, Ritz J *et al*. Methods for pooling results of epidemiologic studies: the Pooling Project of Prospective Studies of Diet and Cancer. *Am J Epidemiol* 2006;**163:**1053–64.

[21] Goldbohm RA, van den Brandt PA, Brants HA *et al*. Validation of a dietary questionnaire used in a large-scale prospective cohort study on diet and cancer. *Eur J Clin Nutr* 1994;**48:**253–65.

[22] Goldbohm RA, van't Veer P, van den Brandt PA *et al*. Reproducibility of a food frequency questionnaire and stability of dietary habits determined from five annually repeated measurements. *Eur J Clin Nutr* 1995;**49:**420–29.

[23] van den Brandt PA, Goldbohm RA, van't Veer P *et al*. A large-scale prospective cohort study on diet and cancer in The Netherlands. *J Clin Epidemiol* 1990;**43:**285–95.

[24] van den Brandt PA, Schouten LJ, Goldbohm RA *et al*. Development of a record linkage protocol for use in the Dutch Cancer Registry for Epidemiological Research. *Int J Epidemiol* 1990;**19:**553–58.

[25] Goldbohm RA, van den Brandt PA, Dorant E. Estimation of the coverage of municipalities by cancer registries and PALGA using hospital discharge data. *Tijdschr Soc Gezondheidsz* 1994;**72:**80–84.

[26] Brink M, de Goeij AF, Weijenberg MP *et al*. K-ras oncogene mutations in sporadic colorectal cancer in The Netherlands Cohort Study. *Carcinogenesis* 2003;**24:**703–10.

[27] Ireland P, Jolley D, Giles GG *et al*. Development of the Melbourne FFQ: a food frequency questionnaire for use in an Australian prospective study involving an ethnically diverse cohort. *Asia Pac J Clin Nutr* 1994;**3:**19–31.

[28] Hughes LA, Schouten LJ, Goldbohm RA *et al*. Self-reported clothing size as a proxy measure for body size. *Epidemiology* 2009;**20:**673–76.

[29] Lohman TG, Roche AF, Martorell R. *Anthropometric Standardization Reference Manual*. Champaign, IL: Kinetics Books, 1988.

[30] Sieben NL, Roemen GM, Oosting J *et al*. Clonal analysis favours a monoclonal origin for serous borderline tumours with peritoneal implants. *J Pathol* 2006;**210:**405–11.

[31] Suraweera N, Duval A, Reperant M *et al*. Evaluation of tumor microsatellite instability using five quasimonomorphic mononucleotide repeats and pentaplex PCR. *Gastroenterology* 2002;**123:**1804–11.

[32] Young J, Barker MA, Simms LA *et al*. Evidence for *BRAF* mutation and variable levels of microsatellite instability in a syndrome of familial colorectal cancer. *Clin Gastroenterol Hepatol* 2005;**3:**254–63.

[33] Lindor NM, Burgart LJ, Leontovich O *et al*. Immunohistochemistry versus microsatellite instability testing in phenotyping colorectal tumors. *J Clin Oncol* 2002;**20:**1043–48.

[34] Korn EL, Graubard BI, Midthune D. Time-to-event analysis of longitudinal follow-up of a survey: choice of the time-scale. *Am J Epidemiol* 1997;**145:**72–80.

[35] Barlow WE, Ichikawa L, Rosner D *et al*. Analysis of case–cohort designs. *J Clin Epidemiol* 1999;**52:**1165–72.

[36] Lunn M, McNeil D. Applying Cox regression to competing risks. *Biometrics* 1995;**51:**524–32.

[37] English DR, Young JP, Simpson JA *et al*. Ethnicity and risk for colorectal cancers showing somatic *BRAF* V600E mutation or CpG island methylator phenotype. *Cancer Epidemiol Biomarkers Prev* 2008;**17:**1774–80.

[38] de Vogel S, Bongaerts BW, Wouters KA *et al*. Associations of dietary methyl donor intake with MLH1 promoter hypermethylation and related molecular phenotypes in sporadic colorectal cancer. *Carcinogenesis* 2008;**29:**1765–73.

[39] de Vogel S, Weijenberg MP, Herman JG *et al*. MGMT and MLH1 promoter methylation versus APC, KRAS and *BRAF* gene mutations in colorectal cancer: indications for distinct pathways and sequence of events. *Ann Oncol* 2009;**20:** 1216–22.

[40] de Vogel S, Wouters KA, Gottschalk RW *et al*. Genetic variants of methyl metabolizing enzymes and epigenetic regulators: associations with promoter CpG island

hypermethylation in colorectal cancer. *Cancer Epidemiol Biomarkers Prev* 2009;**18:**3086–96.
[41] Bolton-Smith C, Woodward M, Tunstall-Pedoe H *et al*. Accuracy of the estimated prevalence of obesity from self reported height and weight in an adult Scottish population. *J Epidemiol Community Health* 2000;**54:**143–48.
[42] Nyholm M, Gullberg B, Merlo J *et al*. The validity of obesity based on self-reported weight and height: implications for population studies. *Obesity;***15:**197–208.
[43] Spencer EA, Appleby PN, Davey GK *et al*. Validity of self-reported height and weight in 4808 EPIC-Oxford participants. *Public Health Nutr* 2002;**5:**561–65.
[44] Wada K, Tamakoshi K, Tsunekawa T *et al*. Validity of self-reported height and weight in a Japanese workplace population. *Int J Obes* 2005;**29:**1093–99.
[45] Gordon I, Boffetta P, Demers PA. A case study comparing a meta-analysis and a pooled analysis of studies of sinonasal cancer among wood workers. *Epidemiology* 1998;**9:**518–24.
[46] Slattery ML, Curtin K, Anderson K *et al*. Associations between cigarette smoking, lifestyle factors, and microsatellite instability in colon tumors. *J Natl Cancer Inst* 2000;**92:** 1831–36.
[47] Smits KM, Cleven AH, Weijenberg MP *et al*. Pharmacoepigenomics in colorectal cancer: a step forward in predicting prognosis and treatment response. *Pharmacogenomics* 2008;**9:**1903–16.
[48] Imai K, Yamamoto H. Carcinogenesis and microsatellite instability: the interrelationship between genetics and epigenetics. *Carcinogenesis* 2008;**29:**673–80.
[49] Hughes LA, Khalid-de Bakker CA, Smits KM *et al*. The CpG island methylator phenotype in colorectal cancer: progress and problems. *Biochim Biophys Acta* 2012;**1825:** 77–85.
[50] Hinoue T, Weisenberger DJ, Lange CP *et al*. Genome-scale analysis of aberrant DNA methylation in colorectal cancer. *Genome Res* 2012;**22:**271–82.
[51] Ogino S, Kawasaki T, Kirkner GJ *et al*. CpG island methylator phenotype-low (CIMP-low) in colorectal cancer: possible associations with male sex and KRAS mutations. *J Mol Diagn* 2006;**8:**582–88.
[52] Yagi K, Akagi K, Hayashi H *et al*. Three DNA methylation epigenotypes in human colorectal cancer. *Clin Cancer Res* 2010;**16:**21–33.
[53] Dirx MJ, van den Brandt PA, Goldbohm RA *et al*. Energy restriction early in life and colon carcinoma risk: results of The Netherlands Cohort Study after 7.3 years of follow-up. *Cancer* 2003;**97:**46–55.
[54] Frankel S, Gunnell DJ, Peters TJ *et al*. Childhood energy intake and adult mortality from cancer: the Boyd Orr Cohort Study. *BMJ* 1998;**316:**499–504.
[55] Hughes LA, van den Brandt PA, Goldbohm RA *et al*. Childhood and adolescent energy restriction and subsequent colorectal cancer risk: results from The Netherlands Cohort Study. *Int J Epidemiol* 2010;**39:**1333–44.
[56] Svensson E, Grotmol T, Hoff G *et al*. Trends in colorectal cancer incidence in Norway by gender and anatomic site: an age-period-cohort analysis. *Eur J Cancer Prev* 2002;**11:** 489–95.
[57] Svensson E, Moller B, Tretli S *et al*. Early life events and later risk of colorectal cancer: age-period-cohort modelling in the Nordic countries and Estonia. *Cancer Causes Control* 2005;**16:**215–23.
[58] Hughes LA, van den Brandt PA, de Bruine AP *et al*. Early life exposure to famine and colorectal cancer risk: a role for epigenetic mechanisms. *PloS One* 2009;**4:**e7951.

Published by Oxford University Press on behalf of the International Epidemiological Association *International Journal of Epidemiology* 2012;**41**:1072–1075

doi:10.1093/ije/dys076

Commentary: Lifestyle factors and colorectal cancer microsatellite instability—molecular pathological epidemiology science, based on unique tumour principle

Shuji Ogino[1,2,3,4]* and Edward Giovannucci[1,4,5,6]

[1]Cancer Epidemiology Program, Dana-Farber/Harvard Cancer Center, [2]Department of Pathology, Brigham and Women's Hospital, and Harvard Medical School, [3]Department of Medical Oncology, Dana-Farber Cancer Institute, [4]Department of Epidemiology, Harvard School of Public Health, [5]Department of Nutrition, Harvard School of Public Health and [6]Channing Laboratory, Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA

*Corresponding author. Center for Molecular Oncologic Pathology, Dana-Farber Cancer Institute, Brigham and Women's Hospital, 450 Brookline Ave, Room JF-215C, Boston, MA 02215, USA.
E-mail: shuji_ogino@dfci.harvard.edu

Accepted 16 April 2012